Rebuttal for Submission665

Catch My Drift: Microscopic Domain Adaptation for Mitigating Non-Stationarity in EEG Signals Manifesting as Temporal Drift

April 11, 2024

Reviewer 4 (Overall score: 3)

Comment 1: *The title might be misleading for researchers focused on drifting in time series. This paper specifically addresses drift in covariances rather than in raw time series. The title should more clearly reflect this to avoid confusion among readers.*

Response:

Thanks for your valuable comment. We would like to convey that our work is indeed concerned with mitigating the effects of non-stationarity on classification methods in time series in general and EEG in particular in our paper. In our evaluation and datasets, we have observed that non-stationarity in EEG time-series manifests itself as drift (which we call temporal drift in our evaluation of a few public datasets) in covariance matrices which represent the raw signals.

To elaborate on this further, we would approach addressing your comment in the following two ways:

- Briefly describing the algorithm that generates these covariance matrices from raw data to show that each covariance matrix captures the spatiotemporal distribution of the raw EEG signals at a given instant.
- Generating another synthetic dataset with non-stationarity and demonstrating how this non-stationarity can be visualized by observing the temporal drift in the resultant SPD covariance matrices.

0.1 xDAWN Algorithm to Generate Covariance Matrices

Rivet et al. [1] proposed the xDAWN algorithm as a robust way to estimate the evoked response of noisy EEG time-series signals. The method estimates the best EEG spatial filters for an evoked response that maximizes the SNR and the raw EEG signals are then projected into the estimated signal subspace. This projection of the raw EEG signals generates covariance matrices where each covariance matrix is representative of the raw EEG time series that generated it. The spatiotemporal location of the covariance matrix on the signal subspace represents the signal characteristics of the corresponding raw EEG signal in the signal space. This observation is then used by various classification algorithms like [2] and [3] to classify EEG time-series data by projecting it on the signal subspace and then classifying the resultant covariance matrices.

The drift in these covariance matrices on a Riemannian manifold is indicative of a gradual transition in the statistical properties of the raw EEG signals. This point can be further demonstrated by the following example where we construct a synthetic dataset with non-stationary statistical properties (a constantly changing mean signal in this instant) and show that this manifests as a temporal drift in the resultant covariance matrices.

0.2 Non-Stationary Synthetic Dataset

We generate a synthetic dataset with non-stationarity to demonstrate how this can manifest as temporal drift in the signal subspace.

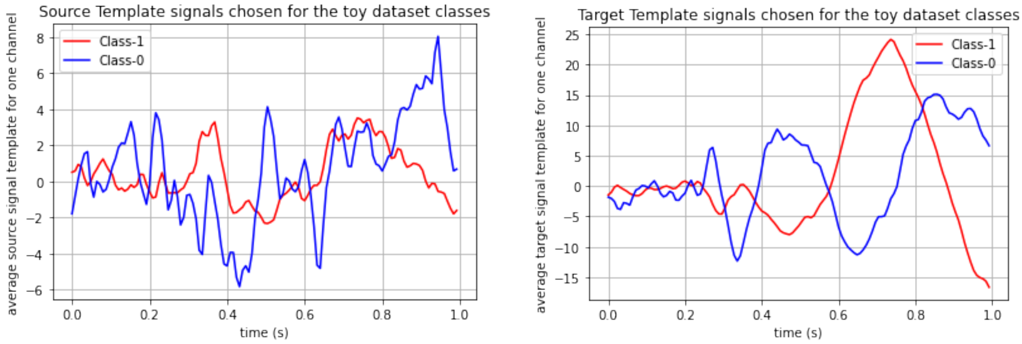
For this, we use two classes of evoked responses, class-0 and class-1. Typically to generate several EEG signal segments corresponding to these classes, we add a zero-mean Gaussian noise vector with a specific variance. However, that results in a stationary EEG process as the mean, variance, and other statistical properties of the resultant time series are stationary (as the time series is a mixture of a fixed EEG template and a stationary noise process). In the paper, we generated such a stationary dataset and showed its stationary properties as well.

For this exercise, in order to generate a non-stationary dataset, we constantly vary the class-0 and class-1 templates on a trial-by-trial basis. We start out by creating two templates for each class (two templates for class-0 and two templates for class-1), thus 4 EEG templates in total. Each signal corresponding to a specific class is then a weighted combination of the two templates and the weighing factor is changed constantly for each sample.

If we denote class-0 templates as C_0^1 and C_0^2 respectively and class-1 templates as C_1^1 and C_1^2, where $C_i^j \in \mathbb{R}^{CXN}$ where C is the number of EEG channels (electrodes) and N is the number of samples in the template, then each EEG time-series signal generated for each class can be represented as:

$$S_0^i = \frac{i}{K} * C_0^1 + \frac{(K-i)}{K} * C_0^1 + N \tag{1}$$

$$S_1^i = \frac{i}{K} * C_1^1 + \frac{(K-i)}{K} * C_1^1 + N \tag{2}$$



(a) Signal templates 1 for synthetic dataset (b) Signal templates 2 for synthetic dataset

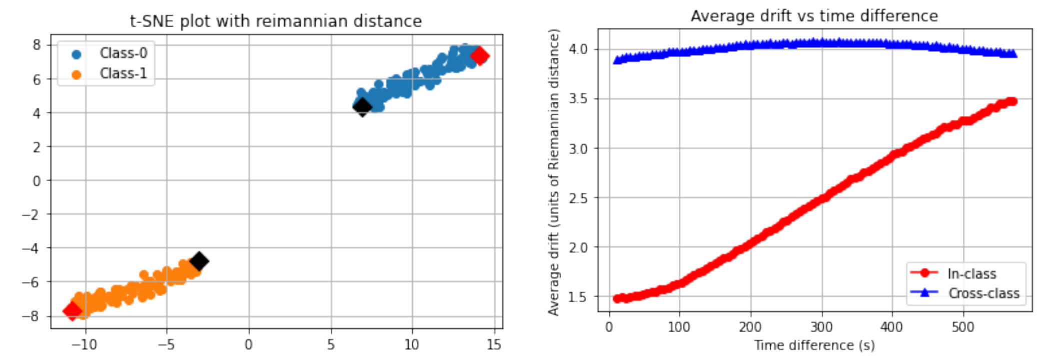
Figure 1: Synthetic dataset using two templates for two classes by adding Gaussian noise

Where K is the total number of signals we generate for each class (in our example $K = 100$), and N is a random noise vector and $N \in \mathbb{R}^{CXN}$. We present the two templates used in Fig 1(a) and 1(b)

We generate K covariance matrices from these K EEG time-series signals using the xDAWN algorithm [1] and plot their signal subspace variation using t-SNE in Fig 2(a). The black-colored diamond-shaped marker denotes the first point for their respective class and the red-colored diamond-shaped point corresponds to the K^{th} signal for their respective class (or 100^th in this instance). We can see the almost linear variation in the spatio-temporal characteristics of the plotted covariance matrices and this is corroborated by Fig 2(b) which shows an increasing in-class temporal drift in the covariance matrices as a result of this non-stationarity. The cross-class temporal drift also varies but its variation is almost negligible compared to that of the in-class drift. Finally, we present the drift distribution per time-difference instant in Fig 0.2 for in-class and cross-class drift, where we see that the temporal distribution of the Riemannian distance between different matrices of the same class changes.

Thus, we have attempted to demonstrate in this document how the non-stationarity present in time-series data can manifest itself as a temporal drift present in the signal covariance matrices which are the projections of the raw signal data on a signal subspace estimated by the xDAWN spatial filtering algorithm.

We generated a synthetic dataset with changing mean template as a function of time (sample number) and demonstrated how this non-stationarity was evident in the temporal drift observed in the signal covariance matrices. By domain-adapting these covariance matrices, we can reduce this temporal drift and in turn mitigate the non-stationarity in our dataset, which in-turn will help in generalization of classification methods.



(a) tSNE graphs of the covariance matrices of the two classes

(b) Average drift as a function of sample time difference

Figure 2: Plots of tSNE for the non-stationary dataset and the corresponding within-class and cross-class temporal drift

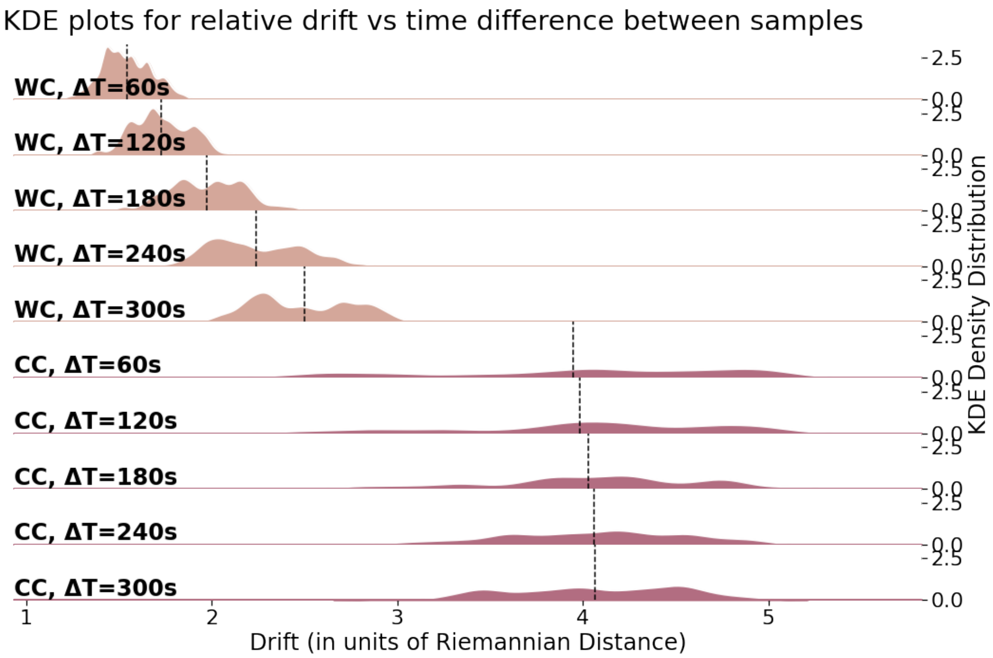


Figure 3: Plot of within-class and cross-class drift distribution

References

[1] B. Rivet*, A. Souloumiac, V. Attina, and G. Gibert, “xdawn algorithm to enhance evoked potentials: Application to brain–computer interface,” *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 8, pp. 2035–2043, 2009.

[2] P. Zanini, M. Congedo, C. Jutten, S. Said, and Y. Berthoumieu, “Transfer learning: A riemannian geometry framework with applications to brain–computer interfaces,” *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 5, pp. 1107–1116, 2018.

[3] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, “Riemannian geometry applied to bci classification,” in *Latent Variable Analysis and Signal Separation*, V. Vigneron, V. Zarzoso, E. Moreau, R. Gribonval, and E. Vincent, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 629–636.